THE POTENTIAL EFFECT OF STATISTICAL DEPENDENCE IN THE ANALYSIS OF DATA IN JURY DISCRIMINATION CASES: *MOULTRIE v. MARTIN* RECONSIDERED

Joseph L. Gastwirth
Weiwen Miao*

ABSTRACT: This article describes a statistical analysis that incorporates the dependence between consecutive grand juries in jurisdictions that allow jurors to serve a second year. This holdover juror system is used in South Carolina and is permitted in California. The appropriate modifications to the usual standard deviation analysis are presented and applied to data from *Moultrie v. Martin*, 690 F.2d 1078 (4th Cir. 1982). The analysis shows that ignoring the dependence made the statistical evidence of discrimination appear stronger than it was. Furthermore, the standard method, which ignores the dependence, also exaggerates the power of the hypothesis test to determine whether there was a potentially discriminatory selection process.

CITATION: Joseph L. Gastwirth and Weiwen Miao, The Potential Effect of Statistical Dependence in the Analysis of Data in Jury Discrimination Cases: *Moultrie v. Martin* Reconsidered, 43 Jurimetrics J. 115–128 (2002).

---

*Joseph L. Gastwirth is Professor of Statistics and Economics at George Washington University. He is a visiting scientist at the Division of Cancer, Epidemiology and Genetics at the National Cancer Institute. Weiwen Miao is Assistant Professor at Macalester College. This research was supported in part by a grant from the National Science Foundation. A version was presented at the Fifth International Conference on Forensic Statistics, Venice, Italy, September 2002. It is a pleasure to thank Leonard E. Klein, Research Librarian at the George Washington University Law School, for his most valuable assistance.

A fair jury-selection process is critical to the American legal system. The Constitution requires that a "fair cross section of the community" be represented on both petit and grand juries and prohibit intentional discrimination on the basis of race.[1] Thus, systematic under-representation of a distinct segment of the community is impermissible,[2] and in *Castaneda v. Partida*,[3] the Supreme Court used statistical hypothesis testing to measure the difference between the racial composition of venire-persons and their fraction of eligible jury candidates in the community.[4]

The standard analysis assumes that the members of the eligible population are chosen independently each year for service on the venire. This assumption is not always true. For example, in *Moultrie v. Martin*,[5] the Fourth Circuit considered a slightly different system used to select grand jurors in South Carolina. Each year, from the population of potential jurors, twelve individuals were chosen at random for service on the grand jury of the county. For the next year, six of these twelve were randomly selected to remain on the grand jury, and twelve new people were added. Thus, a total of eighteen people served on the grand jury each year. Because of this overlapping in service for six panel members, the racial compositions of the yearly grand juries were no longer statistically independent.[6]

This article explores the effect of this dependence on the usual "standard deviation analysis." Such positive dependence changes the distribution of statistics, such as the total number of minority jurors.[7] Assuming independence for the sake of convenience can have serious consequences in a wide variety of applications.[8]

---

1. *See* David H. Kaye, *Statistical Analysis in Jury Discrimination Cases*, 25 JURIMETRICS J. 274, 275 (1985) (citing Taylor v. Louisiana, 419 U.S. 522, 530 (1975)); David H. Kaye, *Statistical Evidence of Discrimination in Jury Selection*, *in* STATISTICAL METHODS IN DISCRIMINATION LITIGATION 15 (David H. Kaye & Mikel Aickin eds., 1986) (citing *Taylor*, 419 U.S. at 530).

2. *See* Duren v. Missouri, 439 U.S. 357, 364 (1979).

3. 430 U.S. 482, 494 (1977).

4. *See* Michael O. Finkelstein, *The Application of Statistical Decision Theory to the Jury Discrimination Cases*, 80 HARV. L. REV. 338 (1966). The data and the method of analysis in *Castaneda* are discussed in 1 JOSEPH L. GASTWIRTH, STATISTICAL REASONING IN LAW AND PUBLIC POLICY 159–67 (1988), and HANS ZEISEL & DAVID H. KAYE, PROVE IT WITH FIGURES: EMPIRICAL METHODS IN LAW AND LITIGATION 177–84 (1997).

5. 690 F.2d 1078, 1080 (4th Cir. 1982).

6. Kaye, *Statistical Analysis in Jury Discrimination Cases*, *supra* note 1, at 285 n.47, noted that this was an important complication but ignored it to follow the opinion).

7. *See* Joseph Gastwirth & Herman Rubin, *Effect of Dependence on the Level of Some One-Sample Tests*, 66 J. AM. STAT. ASS'N 816 (1971).

8. *See* RUPERT G. MILLER, JR., BEYOND ANOVA, BASICS OF APPLIED STATISTICS 32–36, 214–20 (1986); William Kruskal, *Miracles and Statistics: The Casual Assumption of Independence*, 83 J. AM. STAT. ASS'N 929 (1988); *see also* RICHARD A. POSNER, FRONTIERS OF LEGAL THEORY 374–77 (2001) (discussing the role of the assumption of mathematical independence between individual pieces of evidence in models of the process of weighing the totality of the evidence). The problem has arisen, for instance, in risk analyses used in environmental cases. *See* Joseph L. Gastwirth, *The Potential Effect of Unchecked Statistical Assumptions: A Fault in* San Luis Obispo Mothers for Peace v. United States Nuclear Regulatory Commission, 9 J. ENERGY LAW & POL'Y 177 (1989). Although the selection process in *Moultrie* may be uncommon, the holdover system is still

Part I describes the jury selection data in *Moultrie* and the court's analysis. Part II presents formulas that incorporate the dependence between successive grand juries. Part III uses these formulas to reanalyze the *Moultrie* data. We find that ignoring the dependence indicates stronger evidence of discrimination than is the case. Our results also indicate that ignoring the dependence in the racial composition of consecutive juries overestimates the power of the hypothesis test used to detect minority under-representation.

I. THE STATISTICAL ANALYSIS IN *MOULTRIE*

Moultrie v. Martin[9] concerned alleged discrimination against blacks in the selection of grand jurors. The case arose from the murder of a deputy sheriff in Colleton County, South Carolina.[10] A grand jury consisting of 3 blacks and 15 whites indicted a black suspect.[11] At trial, he moved to quash the indictment on the ground that blacks were under-represented on grand juries.[12] The trial court denied the motion, and the defendant was convicted of involuntary manslaughter.[13] He appealed to the federal court arguing that under-representation of blacks on the grand jury violated his constitutional right to a fair trial.[14]

The court of appeals relied on the statistical evidence in making its decision. The binomial model assumes that each juror is chosen *independently* from the population of potential jurors and has probability π of being a minority member. The value π is the minority fraction of the population of eligible jurors determined from external data, such as the latest census figures. Then the total number of minority jurors out of n individuals chosen randomly from the pool of potential jurors has an expected value of $n\pi$ with standard deviation

$$\sqrt{n\pi(1-\pi)}\ .$$

The usual statistic, Z, is based on the difference between the actual number of minority members, A, and its expected value, expressed in terms of the number of standard deviations. Formally,

used in South Carolina. S.C. CODE ANN. § 14-7-1510 (Law Co-op 1976 & Supp. 2001). California maintains a similar law. CAL. PENAL CODE § 901(b) (West 1985 & Supp. 2002). Some advantages of holdover grand jurors are that they speed up the learning process of new jurors and can improve the responsiveness of the agency overseeing the jury system to the recommendations of previous grand juries. See Stephanie A. Doria, Comment, *Adding Bite to the Watchdog's Bark: Reforming the California Civil Grand Jury System*, 28 PAC. L.J. 1115, 1139–44 (1997), for further discussion and references.

9. 690 F.2d 1078 (4th Cir. 1982).
10. *Id.* at 1079.
11. *Id.*
12. *Id.*
13. *Id.*
14. *Id.*

$$Z = \frac{A - n\pi}{\sqrt{n\pi(1-\pi)}}. \tag{1}$$

Z is approximately normally distributed.

The court of appeals determined that blacks comprised 38% of registered voters in 1977.[15] Moreover, the court treated each jury as a random sample of 18 persons from the voter registration lists and hence applied equation (1) with $\pi = 0.38$ and $n = 18$. The data[16] submitted to the court, along with the Z statistics for each year from 1971 to 1977 and the total over all seven years, are given in Table 1. Based on this information, the court decided that the data did not establish a *prima facie* case under the two-to-three-standard-deviation criterion.[17] The opinion observed that while the seven-year total (1971–77) reached -2.9 standard deviations, the total for the six years, 1972–77, only reached -2.0.[18] The court also noted that the analysis made two assumptions: the fraction of minority-registered voters in 1977 was valid for the entire seven-year period, and this value applied to jurors who were held over for a second year.[19]

Table 1. Representation of Blacks on Grand Juries in Colleton County from 1971 to 1977[20]

Year	Actual Number of Blacks	% Blacks	Z Value
1971	1	5.6	-2.84 (-3.4)
1972	5	27.8	-0.89
1973	5	27.8	-0.89
1974	7	38.9	0.08
1975	7	38.9	0.08
1976	4	22.2	-1.38
1977	3	16.7	-1.86
Total	32	25.4	-2.91

15. *Id.* at 1080.

16. The data are given in the opinion, 690 F.2d at 1084 n.11, and in Kaye, *Statistical Analysis in Jury Discrimination Cases, supra* note 1, at 286.

17. 690 F.2d at 1085.

18. *Id.* at 1084.

19. *Id.* at 1084 n.11, 1085. In Part II we will see that the second assumption is correct. The opinion also mentioned that previous tribunals had criticized the petitioner for failing to examine data on the racial composition of individuals exempted for service and repeated its concern over the lack of information on holdover grand jurors. *Id.* at 1085.

20. The Z value for 1971 in the opinion was reported as -3.4. *Id.* at 1084 n.11. However, our calculation shows it equaled -2.84. There were a total of 18 grand jurors in each year.

II. DEPENDENCE BETWEEN THE GRAND JURY COMPOSITIONS IN CONSECUTIVE YEARS

Table 1 was derived under the assumption that each year grand juries were selected as independent random samples from the eligible population. Because of holdovers, however, juries for consecutive years were dependent. To properly calculate the Z statistic when there are holdover jurors, we introduce the following notation:

t: the number of years considered;
π: minority fraction in the population of potential jurors, this fraction is assumed to be constant for all the t years;
n: number of first time jurors;
m: number of holdover jurors;
$X_{1,j}$: number of first time minority jurors in year j, $j = 1,2, \ldots, t$;
$X_{2,j}$: number of holdover minority jurors in year j, $j = 1,2, \ldots, t$;
B_j: number of minority jurors for year j, $B_j = X_{1,j} + X_{2,j}$;
S_t: total number of minority jurors for t years, $S_t = B_1 + B_2 + \ldots + B_t$;
$E(S_t)$: expected value of S_t;
$SD(S_t)$: standard deviation of S_t;
A: the total number of minority jurors who served during the t years.[21]

If there were no discrimination in the selection process, new jurors would be randomly selected each year from the population of potential jurors. Hence, the $X_{1,j}$s are all independent binomial random variables with parameters n and π. Furthermore, the m holdover jurors should be a random sample from the n jurors who served in the *previous* year. Thus, the conditional distribution of $X_{2,j}$ given $X_{1,j-1}$ is hypergeometric. Appendix A shows that $X_{2,j}$ also has a binomial distribution with parameters m and π.[22] Moreover, $X_{2,j}$ only depends on the racial decomposition of year $j-1$, so $X_{1,j}$ and $X_{2,j}$ are independent. Consequently, B_j also has a binomial distribution with parameters $n+m$ and π.[23]

Since S_t is the sum of *dependent* binomial variables, its variance is the sum of the variances of the individual B_js plus the sum of their covariances. Appendix B shows that the covariance between the number of minorities in consecutive years B_j and B_{j+1}, equals $m\pi(1-\pi)$.[24] The covariance between the number of minorities in nonconsecutive years, i.e., B_j and B_{j+s}, for $s > 1$, is 0. Thus, the mean and standard deviation of S_t are given by

$$\mathrm{E}(S_t) = (n+m)t\pi$$

21. To follow the usual "standard deviation" analysis, we need to show that S_t is approximately normally distributed. Then we can express the difference between A and $E(S_t)$ in terms of the number of standard deviations of S. Since B_j and B_{j+1} are dependent due to the holdover jurors, S_t no longer is a sum of independent random variables.

22. This proves that the holdover jurors also had the same minority fraction as the first-time jurors, and hence the court's second assumption was correct. 690 F.2d at 1084 n.11.

23. *See* RICHARD J. LARSEN & MORRIS L. MARX, AN INTRODUCTION TO MATHEMATICAL STATISTICS AND ITS APPLICATIONS 241 (3d ed. 2001).

24. *See infra* Appendix, part B.

and
$$\mathrm{SD}\ (S_t) = \sqrt{[(3t-2)m + tn]\pi(1-\pi)}. \tag{2}$$

The exact distribution of S_t is quite complicated.[25] However, the Central Limit Theorem for dependent random variables[26] implies that S_t is asymptotically normally distributed. When the normal approximation to the binomial distribution is reliable, the standard deviation analysis is applicable. In general, this approximation is quite accurate when the continuity correction is applied and the following two inequalities hold: [27]

$$(n+m)t\pi - 3\sqrt{[(3t-2)m+tn]\pi(1-\pi)} > 0, \tag{3}$$

$$(n+m)t\pi + 3\sqrt{[(3t-2)m+tn]\pi(1-\pi)} < (n+m)t. \tag{4}$$

In other words, when (3) and (4) hold, the Z statistic that incorporates the dependence due to holdovers is[28]

$$Z \approx \frac{(A-(n+m)t\pi)+0.5}{\sqrt{[(3t-2)m+tn]\pi(1-\pi)}}. \tag{5}$$

The corresponding p-value of the test is obtained from a standard normal distribution table.

III. A REANALYSIS OF THE DATA

A. Z Statistics and Corresponding *P*-values for *Moultrie*

In *Moultrie*, $n = 12$, $m = 6$, $\pi = 0.38$, and conditions (3) and (4) hold for any $t \geq 1$. Table 2 reports the results of applying equation (5) to the data in Table 1. Both the Z statistics and the corresponding p-values are given. For example, for the period 1971–76, $t = 7$ years, and the actual number of black jurors for this 7-year period is $A = 32$. Hence, according to (5),

$$Z = \frac{(32-(12+6)\times 7\times 0.38)+0.5}{\sqrt{[(3\times 7-2)\times 6+7\times 12]\times 0.38\times(1-0.38)}} = -2.252.$$

25. *See infra* Appendix, part B.

26. *See, e.g.*, Wassily Hoeffding & Herbert Robbins, *The Central Limit Theorem for Dependent Random Variables*, 15 DUKE MATHEMATICAL J. 773 (1948); DONALD A.S. FRASER, NONPARAMETRIC METHODS IN STATISTICS 215–23 (1957).

27. *See* LARSEN & MARX, *supra* note 23, at 268.

28. The 0.5 in equation (5) is the continuity correction that increases the accuracy of the normal approximation. *See* LARSEN & MARX, *supra* note 23, at 268.

For comparative purposes, the results assuming independence between consecutive years are also given in Table 2.[29]

Table 2. Reanalysis of the Data Incorporating the Dependence[30]

Time period	Number of Blacks	Z value Ind	Z value Dep	p-value Ind	p-value Dep
1976–1977	7	-2.122	-1.838	0.034	0.066
1975–1977	14	-1.688	-1.404	0.091	0.160
1974–1977	21	-1.423	-1.162	0.155	0.245
1973–1977	26	-1.672	-1.350	0.094	0.177
1972–1977	31	-1.891	-1.516	0.059	0.129
1971–1977	32	-2.823	-2.252	0.005	0.024

Comparing the Z statistics in Table 2, it is clear that when the dependence is ignored, their magnitudes are inflated. Thus, assuming independence between consecutive juries improperly increases the statistical significance (reduces the p-value) of the Z statistic. More importantly, as the case was brought in 1977, for the most recent two-year period (1976–77), incorporating dependence yields a Z value of -1.84, while assuming independence gives a Z value of -2.12. Had the court applied the two-standard-deviation criterion and accounted for the dependence, it would have found that no *prima facie* case was established. Instead, the court observed that the 1971 data appeared considerably different from the other years and decided to disregard it.[31] For the 1972–77 period, the value of the Z statistic (-1.89) was calculated assuming that juries in consecutive years were independent. The corresponding p-value of 0.059 is close to significance. The Z statistic incorporating dependence is -1.516, which is not significant even at the 10% level (p-value $= 0.129$). This Z statistic of -1.516 is consistent with the other Z statistics for time periods not including 1971, if the two-standard-deviation threshold for the establishment of a *prima facie* case of discrimination is used. On the other hand, when combining all seven years, the analysis incorporating dependence results in a Z value of -2.25. Though the

29. The Z values for the independent case were calculated by using

$$Z = \frac{(A-(n+m)t\pi)+0.5}{\sqrt{(n+m)t\pi(1-\pi)}}$$

to incorporate the continuity correction. Thus, these values differ slightly from those in Table 1.

30. The "Dep" columns present the calculations incorporating the dependence between consecutive years, while the "Ind" columns give the results assuming independence. All the p-values are two-sided.

31. 690 F.2d at 1084. The appropriateness of the court's deeming that 1971 was a statistical outlier due to an unusual situation might be questioned. Legally, one would wish to examine other cases to decide how many prior years of data are typically considered in jury discrimination cases. Statistically, the issue is whether the value of $\pi = 0.38$ determined from 1977 accurately represented the black fraction of eligible jurors in 1971.

magnitude of this Z statistic is noticeably less than the value of -2.82 obtained assuming independence, it still exceeds the two-standard-deviations threshold.

This indicates that determining how many previous years of data should be examined and how one should handle data for a year that differs dramatically from the other years can be quite important. In *Moultrie*, apparently neither party presented evidence that might explain what happened in 1971.[32]

If there were no dependence, the standard deviation of S_t would be

$$\mathrm{SD}(S_t) = \sqrt{(n+m)t\pi(1-\pi)}. \tag{6}$$

This is smaller than the correct one given in (2) if $t > 1$. In general, whenever the yearly data are positively correlated, as in *Moultrie*, assuming independence will underestimate the standard deviation. This yields a p-value smaller than it should be, making the statistical evidence of discrimination appear stronger than it is.

B. Dependence and Power

The underestimate of the standard deviation also causes the power of the test :o be overestimated. Let π_0 and π_1 be the minority fractions, and let σ_0 and σ_1 be :he standard deviations of S_t, under the null and alternative hypotheses, respec- :ively.[33] Let $Z_{\alpha/2}$ be the upper $\alpha/2$ percentile of the standard normal distribution, .e., $Z_{\alpha/2}$ is the value such that $P(Z > Z_{\alpha/2}) = \alpha/2$. For the two-tailed test, the power s

$$Power = P\left(Z \le \frac{(n+m)t(\pi_0 - \pi_1)}{\sigma_1} - z_{\alpha/2}\frac{\sigma_0}{\sigma_1}\right) + P\left(Z \ge \frac{(n+m)t(\pi_0 - \pi_1)}{\sigma_1} + z_{\alpha/2}\frac{\sigma_0}{\sigma_1}\right) \tag{7}$$

Applying formula (7) to *Moultrie* for $t = 7$ years and $\alpha = 0.05$ yields the ollowing table:[34]

---

32. Conceivably, the state might have shown that the minority population of the county ıcreased dramatically at the end of 1971 or beginning of 1972 so that blacks formed substantially :ss than 38% of those eligible for service.

33. Using equation (2), which incorporates dependence properly, gives the following standard eviations:

$$\sigma_0 = \sqrt{[(3t-2)m + tn]\pi_0(1-\pi_0)}$$

ıd

$$\sigma_1 = \sqrt{[(3t-2)m+tn]\pi_1(1-\pi_1)}\cdot$$

n the other hand, ignoring dependence, equation (6) gives standard deviations of:

$$\sigma_0 = \sqrt{(n+m)t\pi_0(1-\pi_0)}$$

ıd

$$\sigma_1 = \sqrt{(n+m)t\pi_1(1-\pi_1)}.$$

34. To illustrate the use of equation (7), consider the fifth row in table 3: $\pi_1 = 0.3$. Notice that $= 12$, $m = 6$, and $\pi_0 = 0.38$. For the test with $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$. Thus, incorporating pendence, the standard deviations are

Table 3. Power of a Two-Tailed Test for $\pi = 0.38$ for 1971–77 Period[35]

% Black (true π)	Power-Ind	Power-Dep	Difference
0.10	1.000	1.000	0.000
0.15	1.000	0.999	0.001
0.20	0.996	0.951	0.046
0.25	0.880	0.688	0.191
0.30	0.454	0.304	0.150
0.35	0.102	0.081	0.021
0.40	0.077	0.068	0.009
0.45	0.370	0.258	0.112
0.50	0.786	0.597	0.188

From Table 2, we see that ignoring dependence can overestimate the power of the test by as much as 19%. By analogy with the "four-fifths" rule,[36] one expects that courts would be interested in detecting minority under-representation of 20% or more below their fraction of eligible jurors. For *Moultrie*, this implies that a minority fraction of grand jurors of $0.8 \times 0.38 = 0.304$ or less is legally meaningful. Table 3 shows that for the *Moultrie* data, the greatest loss of power

---

$$\sigma_0 = \sqrt{[(3\times7-2)\times6+7\times12]\times0.38\times(1-0.38)}$$

and

$$\sigma_1 = \sqrt{[(3\times7-2)\times6+7\times12]\times0.3\times(1-0.3)}$$

and the power is:

$$P\left(Z \le \frac{(12+6)(7)(0.38-0.3)}{6.448} - 1.96\times\frac{6.830}{6.448}\right) + P\left(Z \ge \frac{(12+6)(7)(0.38-0.3)}{6.448} + 1.96\times\frac{6.830}{6.448}\right)$$

$$= P(Z \le -0.513) + P(Z \ge 3.639) = 0.304.$$

On the other hand, if the yearly data were independent, then the standard deviations would be

$$\sigma_0 = \sqrt{(12+6)(7)(0.38)(1-0.38)} = 5.448$$

and

$$\sigma_1 = \sqrt{(12+6)(7)(0.3)(1-0.3)} = 5.144.$$

Consequently, the power is:

$$P\left(Z \le \frac{(12+6)(7)(0.38-0.3)}{5.144} - 1.96\times\frac{5.448}{5.144}\right) + P\left(Z \ge \frac{(12+6)(7)(0.38-0.3)}{6.448} + 1.96\times\frac{5.448}{5.144}\right)$$

$$= P(Z \le -0.116) + P(Z \ge 4.036) = 0.454.$$

35. The table presents the power of the test that the black fraction of grand jurors was consistent with their eligibility fraction, 0.38, i.e., $\pi = 0.38$ against various alternative values of π, using S_t as the test statistic. The significance level is $\alpha = 0.05$.

36. Adverse Impact and the "Four-Fifths Rule," 29 C.F.R. § 1607.4 (D) (1998). See MICHAEL O. FINKELSTEIN & BRUCE LEVIN, STATISTICS FOR LAWYERS 39 (2d ed. 2001), for a discussion and illustration of the rule.

occurs when one is trying to detect minority fractions in the important range of 0.25 to 0.30.

Ignoring the statistical dependence between consecutive grand jurors in *Moultrie v. Martin* led to an apparent increase of about one-half a standard deviation unit. In light of the two-to-three-standard-deviations criterion used in discrimination cases, an error of this magnitude can change a nonsignificant result to a significant one. Indeed, for the most recent 1976–77 time period, which might be considered more relevant to a charge filed in 1977, Table 2 shows that incorporating the dependence properly yielded a nonsignificant result of −1.84 standard deviations. Neglecting the dependence leads to a significant result of −2.12 standard deviations. The positive correlation, however, also decreases the power of the statistical test to detect discrimination,[37] which may diminish the importance of a finding that is not statistically significant. For example, if one desires to detect a county that discriminated by limiting a minority group forming 38% of those eligible to serve to only 30% of grand jurors, line five of Table 3 shows that the probability of detecting the discriminatory process is only 0.304.

It is important to ensure that the assumptions underlying a statistical analysis are satisfied—particularly when the statistical evidence is near the cut-off point for significance.

---

37. The level of significance is typically set at 0.05, corresponding to the two-standard-deviation criterion. The effect of positive correlation, however, applies to any pre-set level of significance.

APPENDIX

This Appendix presents several mathematical results referred to in Part II. First, we prove that $X_{2,j}$ has a binomial distribution. Then we show that S_t is asymptotically normally distributed with mean and standard deviation given in (2). Finally we provide the approximate and exact distribution of S_t.

A. The Distribution of $X_{2,j}$ Is Binomial with Parameters *m* and π

PROOF. Since $X_{1,j-1}$ has binomial distribution and the conditional distribution of $X_{2,j}$ given $X_{1,j-1}$ is hypergeometric, for any $i \le m$,

$$P(X_{2,j}=i)=\sum_{k=i}^{n}P(X_{1,j-1}=k)\cdot P(X_{2,j}=i\mid X_{1,j-1}=k)$$

$$=\sum_{k=i}^{n}\binom{n}{k}\pi^k(1-\pi)^{n-k}\frac{\binom{k}{i}\cdot\binom{n-k}{m-i}}{\binom{n}{m}}$$

$$=\binom{m}{i}\sum_{k=i}^{n}\binom{n-m}{k-i}\pi^k(1-\pi)^{n-k}$$

$$=\binom{m}{i}\pi^i(1-\pi)^{m-i}\left[\sum_{k=i}^{n}\binom{n-m}{k-i}\pi^{k-i}(1-\pi)^{n-m-(k-i)}\right]$$

$$=\binom{m}{i}\pi^i(1-\pi)^{m-i}\left[\sum_{l=0}^{n-i}\binom{n-m}{l}\pi^l(1-\pi)^{n-m-l}\right]$$

Notice that $i \le m$, $n-i \ge n-m$ and $\binom{n-m}{l}=0$, when $l > n-m$. So,

$$\sum_{l=0}^{n-i}\binom{n-m}{l}\pi^l(1-\pi)^{n-m-l}=\sum_{l=0}^{n-m}\binom{n-m}{l}\pi^l(1-\pi)^{n-m-l}=1$$

Consequently,

$$P(X_{2,j}=i)=\binom{m}{i}\pi^i(1-\pi)^{m-i}.$$

B. The Mean and Standard Deviation of S_t Under Dependence

Recall for any *j*, B_j has binomial distribution with parameters $(n+m)$ and π, and B_j and B_{j+1} are dependent. Hence S_t is the sum of dependent binomial variables with mean $(n+m)t\pi$, i.e., $E(S_t)=(n+m)t\pi$.

The variance of each B_j is $(n+m)\pi(1-\pi)$. To obtain the variance of S_t, we need to find the covariance among the B_js. Notice that there are no common holdover jurors between the nonconsecutive years, the covariance between those years is 0: $\mathrm{Cov}(B_j, B_{j+s}) = 0$, for $s > 1$. Since the conditional distribution of $X_{2,j}$ given $X_{1,j-1}$ is hypergeometric,

$$E(X_{1,j-1}X_{2,j}) = E\left[E(X_{1,j-1}X_{2,j} \mid X_{1,j-1})\right] = E\left[X_{1,j-1}E(X_{2,j} \mid X_{1,j-1})\right] = E\left[\frac{m}{n}X_{1,j-1}^2\right]$$

$$= \frac{m}{n}\left[n^2\pi^2 + n\pi(1-\pi)\right] = mn\pi^2 + m\pi(1-\pi)$$

ınd

$$\mathrm{Cov}(X_{1,j-1}, X_{2,j}) = \mathrm{E}\left(X_{1,j-1}, X_{2,j}\right) - \mathrm{E}(X_{1,j-1})\mathrm{E}(X_{2,j})$$

$$= mn\pi^2 + m\pi(1-\pi) - n\pi m\pi = m\pi(1-\pi).$$

Notice that $X_{1,j-1}$ is independent of $X_{1,j}$. Also, $X_{2,j-1}$ is independent of $X_{1,j}$ and $X_{2,j}$. Hence,

$$\mathrm{Cov}(B_{j-1}, B_j) = \mathrm{Cov}(X_{1,j-1} + X_{2,j-1}, X_{1,j} + X_{2,j})$$

$$= \mathrm{Cov}(X_{1,j-1}, X_{1,j}) + \mathrm{Cov}(X_{1,j-1}, X_{2,j}) + \mathrm{Cov}(X_{2,j-1}, X_{1,j}) + \mathrm{Cov}(X_{2,j-1}, X_{2,j})$$

$$= \mathrm{Cov}(X_{1,j-1}, X_{2,j}) = m\pi(1-\pi)$$

'his leads to

$$\mathrm{Var}(S_t) = \mathrm{Var}(B_1 + B_2 + \ldots + B_t) = \sum_{j=1}^{t}\mathrm{Var}(B_j) + 2\sum_{i>j}\mathrm{Cov}(B_i, B_j)$$

$$= t(n+m)\pi(1-\pi) + 2(t-1)m\pi(1-\pi)$$

$$= \left[(3t-2)m + tn\right]\pi(1-\pi).$$

onsequently,

$$\mathrm{SD}(S_t) = \sqrt{\left[(3t-2)m + tn\right]\pi(1-\pi)}.$$

C. Approximate and Exact Distribution of S_t

$S_t = B_1 + B_2 + \ldots + B_t$, where the B_js are identically distributed and 1-dependent.[38] By the Central Limit Theorem,[39]

$$\frac{S_t - t(m+n)\pi}{\sqrt{[(3t-2)m+tn]\pi(1-\pi)}}$$

converges to standard normal distribution, as t increases. Consequently, using the continuity correction, for any x,

$$P(S_t \le x) \approx P\left(Z \le \frac{x - [t(m+n)\pi] + 0.5}{\sqrt{[(3t-2)m+tn]\pi(1-\pi)}}\right) \tag{8}$$

provided that t and π satisfy the conditions given in (3) and (4).

To obtain the exact distribution of S_t, we define the new random variables $Y_j = X_{1,j} + X_{2,j+1}$, $j = 1, 2, \ldots, t\text{-}1$. Notice that those Y_j are independent and identically distributed. Each Y_j is the sum of two dependent random variables. Let $a_i = P(Y_j = i)$. Then,

$$a_i = P(Y_j = i) = \sum_{k=0}^{i} P(X_{2,j+1} = i-k \mid X_{1,j} = k) \cdot P(X_{1,j} = k)$$

$$= \sum_{k \ge i/2}^{i} \frac{\binom{k}{i-k}\binom{n-k}{m-i+k}}{\binom{n}{m}} \binom{n}{k} \pi^k (1-\pi)^{n-k}$$

$$= \sum_{k \ge i/2}^{i} \binom{m}{i-k}\binom{n-m}{2k-i} \pi^k (1-\pi)^{n-k}.$$

Notice that

$$S_t = B_1 + B_2 + \ldots + B_t = (X_{1,1} + X_{2,1}) + (X_{1,2} + X_{2,2}) + \ldots + (X_{1,t} + X_{2,t})$$

$$= (X_{1,1} + X_{2,2}) + (X_{1,2} + X_{2,3}) + \ldots + (X_{1,t-1} + X_{2,t}) + (X_{2,1} + X_{1,t})$$

$$= Y_1 + Y_2 + \ldots + Y_{t-1} + (X_{2,1} + X_{1,t})$$

---

38. A sequence of variables $X_1, X_2, \ldots$ is m-dependent if $(X_1, \ldots, X_r)$ is independent of $(X_s, X_{s+1}, \ldots)$ provided $s\text{-}r$ is greater than m. FRASER, *supra* note 26, at 215. In other words, if m or more consecutive X's are removed, the two remaining sets of random variables are independent. In our situation, $m = 1$ as only consecutive years have a common subset of jurors.

39. *See supra* note 26. The asymptotic variance of S_t given by the Central Limit Theorem is $(n+3m)tp(1-p)$. We use the exact variance of S_t given in Appendix, Part B to achieve a better approximation.

Since $X_{2,1}$ and $X_{1,t}$ are binomial random variables, independent of each other, $X_{2,1} + X_{1,t}$ is also binomial with parameters $(n+m)$ and π. Moreover, $X_{2,1} + X_{1,t}$ is independent of Y_j. Together with the mutual independence of the Y_j, $j = 1, \dots, t-1$, we have:

$$P(S_t \le x) = \sum_{i=0}^{x} P\left(Y_1 + Y_2 + \dots + Y_{t-1} + (X_{2,1} + X_{1,t}) = i\right)$$

$$= \sum_{i=0}^{x} \sum_{k \le i} P(Y_1 + Y_2 + \dots + Y_{t-1} = i-k) \cdot P(X_{2,1} + X_{1,t} = k)$$

$$= \sum_{i=0}^{x} \sum_{k \le i} \left[P(X_{2,1} + X_{1,t} = k) \cdot \sum_{b_1+b_2+\dots+b_{t-1}=i-k} P(Y_1 = b_1) \cdot P(Y_2 = b_2) \dots P(Y_{t-1} = b_{t-1}) \right]$$

$$= \sum_{i=0}^{x} \sum_{k=\max(0,i-(t-1)(m+n))}^{\min(i,m+n)} \left[\binom{m+n}{k} \pi^k (1-\pi)^{(m+n)-k} \cdot \sum_{b_1+\dots+b_{t-1}=i-k} \left(a_{b_1} a_{b_2} \dots a_{b_{t-1}}\right) \right] \qquad (9)$$

For *Moultrie v. Martin*, $m = 6$, $n = 12$, and $\pi = 0.38$. Using the above formula, the exact probability of obtaining 32 or fewer minority jurors for the seven-year period is 0.01056. The exact p-value for the two-tailed test is 0.02266.[40] The normal approximation, .024, is sufficiently close for practical use.

It is easy to check that when $p = 0.38$, conditions (3) and (4) are satisfied when $t \ge 1$. For $t = 2$ years, comparing the exact and normal approximation to the probability $P(S_t \le x)$, using equations (9) and (8) respectively, we found that the largest difference is about 0.007. However, when either (3) or (4) fails, the normal approximation should not be used.[41]

40. Since the exact distribution of S_t is not perfectly symmetric, the two-tailed p-value is not twice the probability of 0.01056, the one-tailed p-value. The two-tailed p-value is obtained as follows: the expected number of minority jurors is $7 \times 18 \times 0.38 = 47.88$. Hence the difference between the expected and observed numbers of minority jurors is: $47.88-32 = 15.88$. The two-tailed p-value is $P(S_7 \le 32) + P(S_7 \ge 47.88 + 15.88) = P(S_7 \le 32) + P(S_7 \ge 64) \approx 0.02266$.

41. For example, when $\pi = 0.1$, and $t = 2$ years, the term on the left-hand side of inequality (3) should be positive. However, it equals

$$(12+6)(2)(0.1) - 3\sqrt{[(3(2)-2)(6)+(2)(12)](0.1)(1-0.1)} = -2.64 < 0$$

which means that the condition (3) fails for $t = 2$. Our calculation shows that the difference between the exact and normal approximation to the probability $P(S_2 \le x)$ can be as much as 0.04, which may change one's conclusion based on the 0.05 cutoff. When $t = 2$, the total number of jurors is 30 plus holdovers, a relatively small sample. The accuracy of the normal approximation to the binomial distribution in samples of small and moderate sizes is still an area of active research. *See* Lawrence D. Brown et al., *Interval Estimation for a Binomial Proportion*, 16 STAT. SCI. 101 (2001).